August 3–7, 2025
KDD2025

Characterizing the Risk of Atrial Fibrillation in Cardiac Patients with Exceptional Electrocardiogram Phenotypes

Lieke van den Biggelaar[1], Rianne M. Schouten[1], Ashley de Bie[2], R. Arthur Bouwman[2], Wouter Duivesteijn[1]

[1] Eindhoven University of Technology, Eindhoven, the Netherlands
[2] Catharina Hospital, Eindhoven, the Netherlands

Medical Problem

Atrial Fibrillation Risk after Cardiac Surgery

Patient undergoes cardiac surgery

Medical Problem

Atrial Fibrillation Risk after Cardiac Surgery

Patient undergoes cardiac surgery

100%

Patient recovers at the intensive care unit

Medical Problem

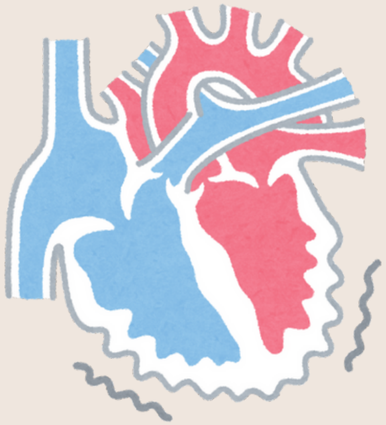
Atrial Fibrillation Risk after Cardiac Surgery

Patient undergoes cardiac surgery

100%

Patient recovers at the intensive care unit

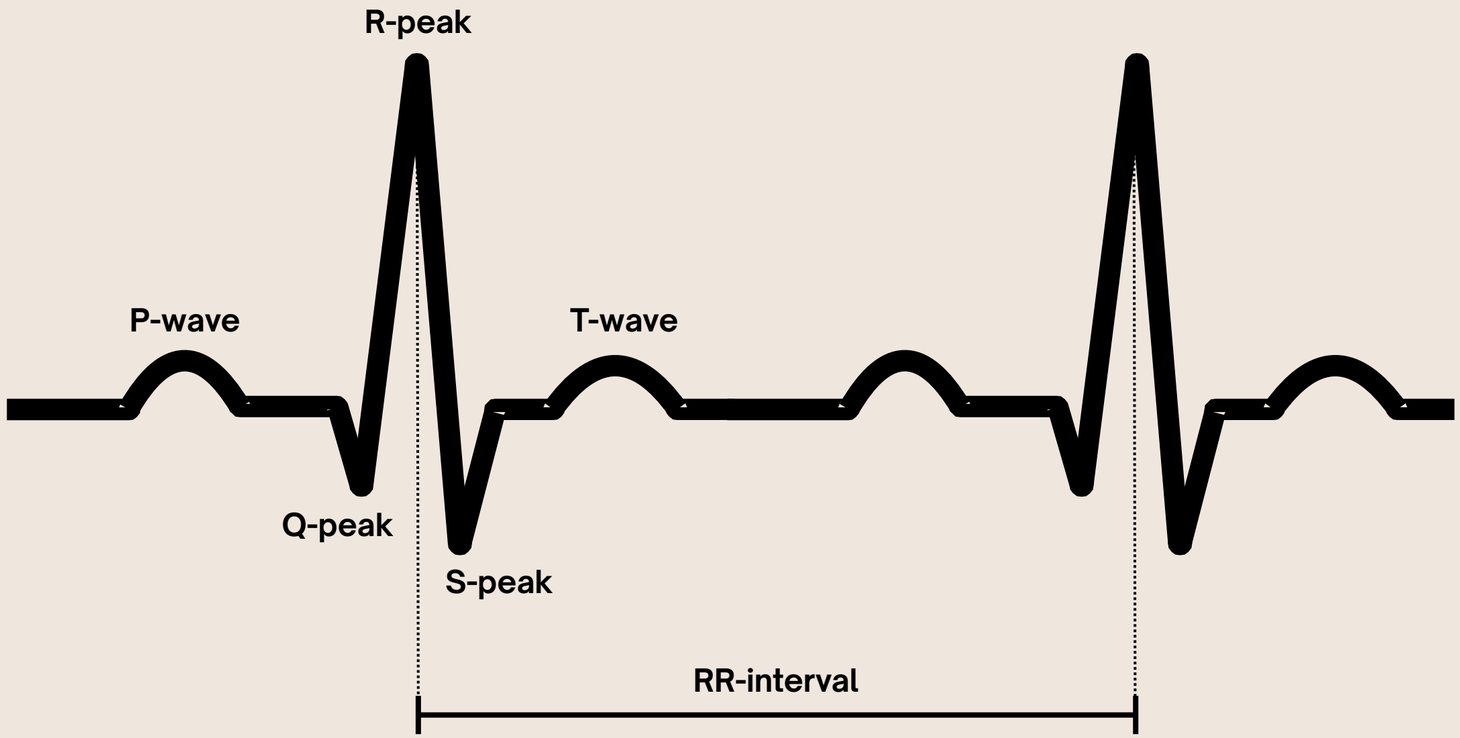
20%-50%



Patient develops atrial fibrillation

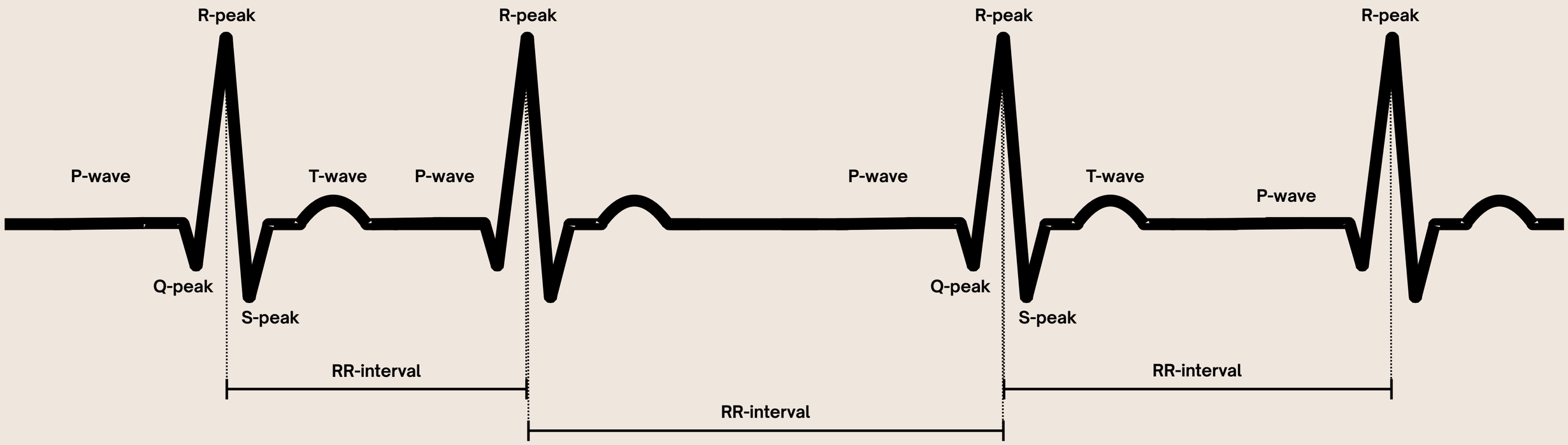
Medical Problem

Current Practice on Atrial Fibrillation Detection



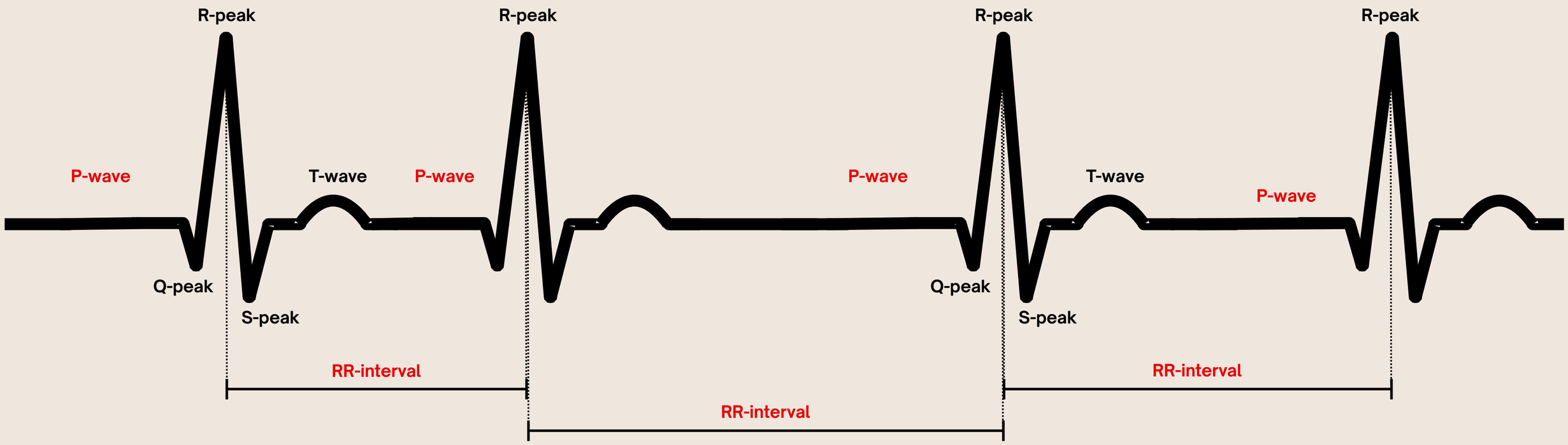
Medical Problem

Current Practice on Atrial Fibrillation Detection



Medical Problem

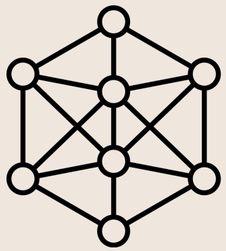
Current Practice on Atrial Fibrillation Detection

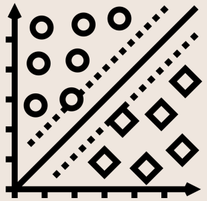


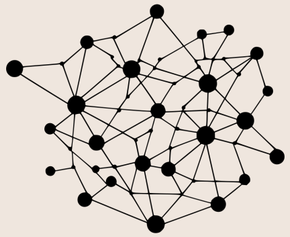
Local Pattern Mining

Automated AF Detection and Prediction

Predictive Methods



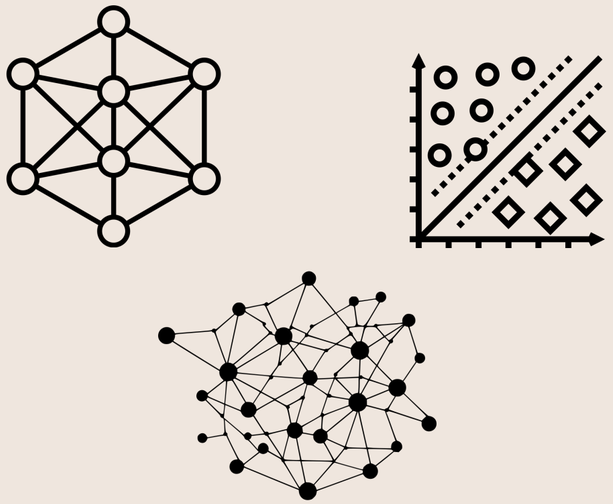




Local Pattern Mining

Automated AF Detection and Prediction

Predictive Methods

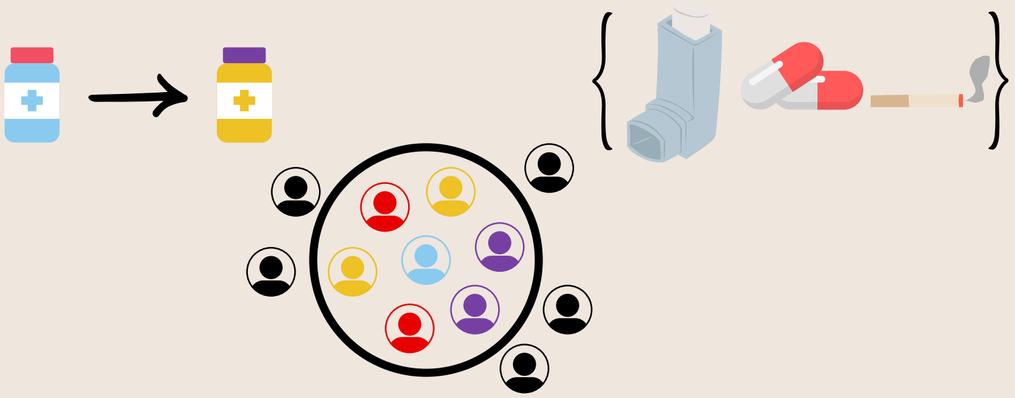


- Global, one-size-fits-all manner
- Difficult to understand for non-technical users
- Complicated deployment in real-world practice

Local Pattern Mining

Automated AF Detection and Prediction

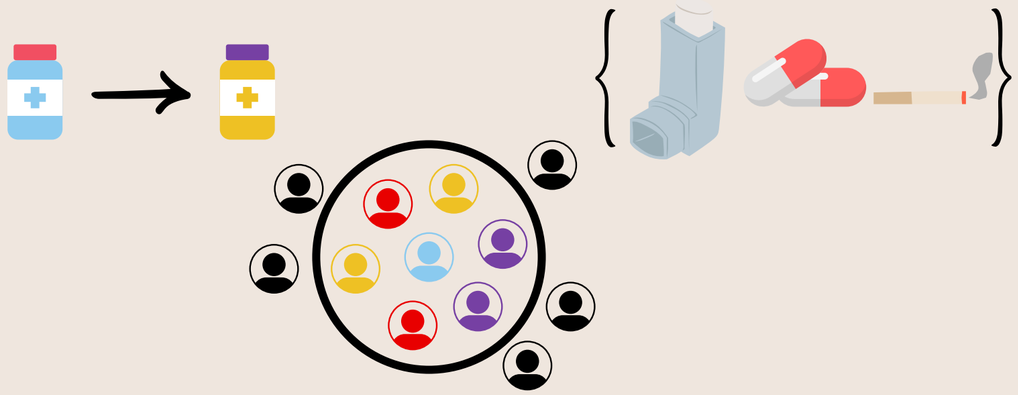
Descriptive Methods



Local Pattern Mining

Automated AF Detection and Prediction

Descriptive Methods

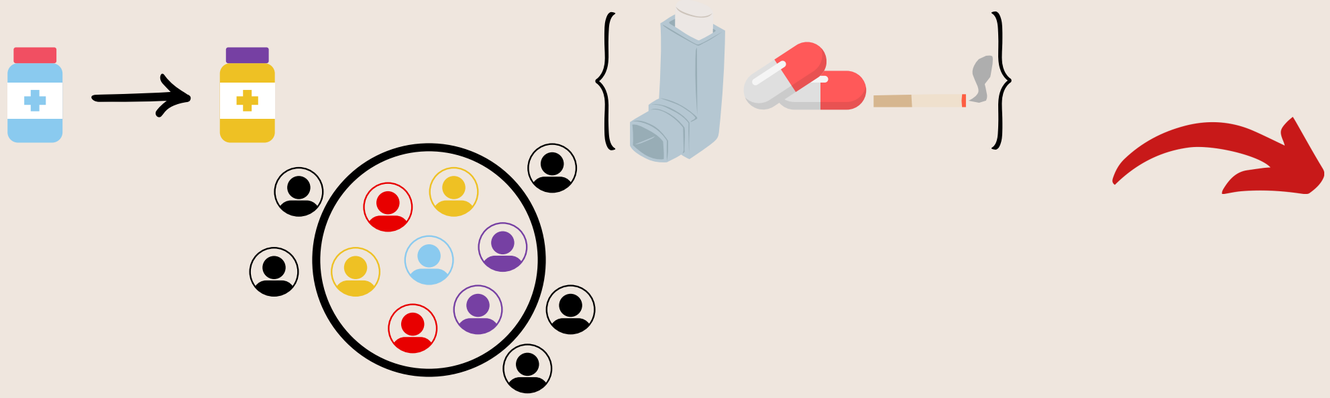


- Structured search for local patterns
- Transparent frameworks
- A move towards stratified medicine
- Automated discovery of subgroups with exceptional AF risks

Local Pattern Mining

Automated AF Detection and Prediction

Descriptive Methods



Local Pattern Mining

- Structured search for local patterns
- Transparent frameworks
- A move towards stratified medicine
- Automated discovery of subgroups with exceptional AF risks

Local Pattern Mining

Exceptional Model Mining

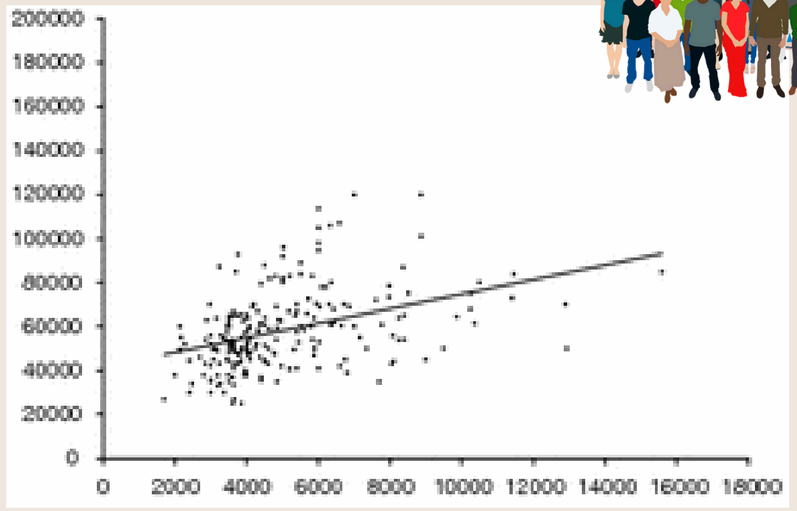
Population

VERSUS

Subgroup

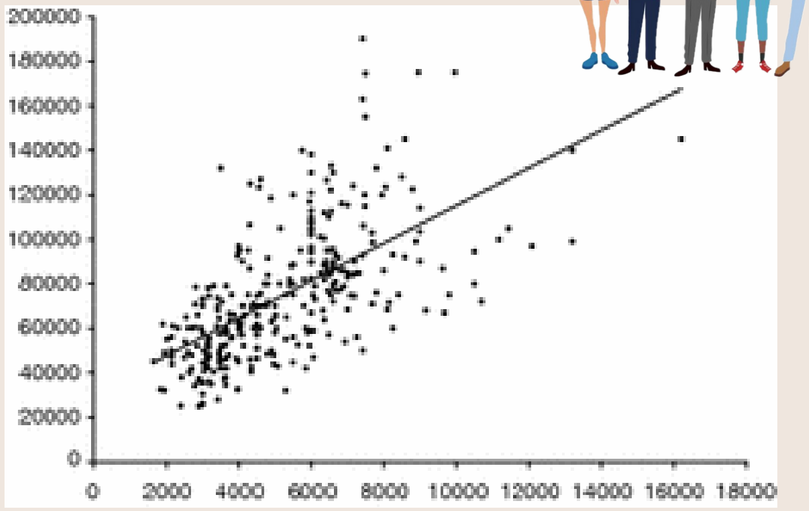
Local Pattern Mining

Exceptional Model Mining



Population

VERSUS



Subgroup

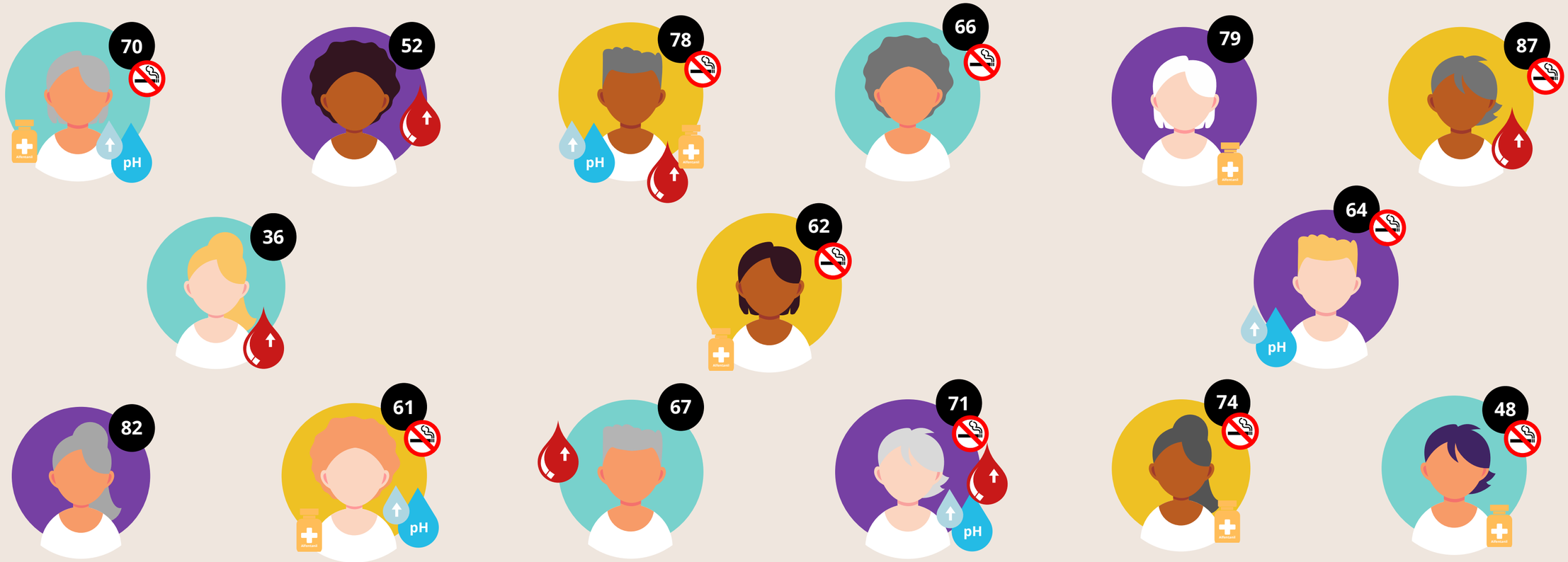
Duivesteijn, W., Feelders, A.J. & Knobbe, A. Exceptional Model Mining. Data Min Knowl Disc30, 47–98 (2016). https://doi.org/10.1007/s10618-015-0403-4

Local Pattern Mining

Exceptional Model Mining for Stratified Medicine

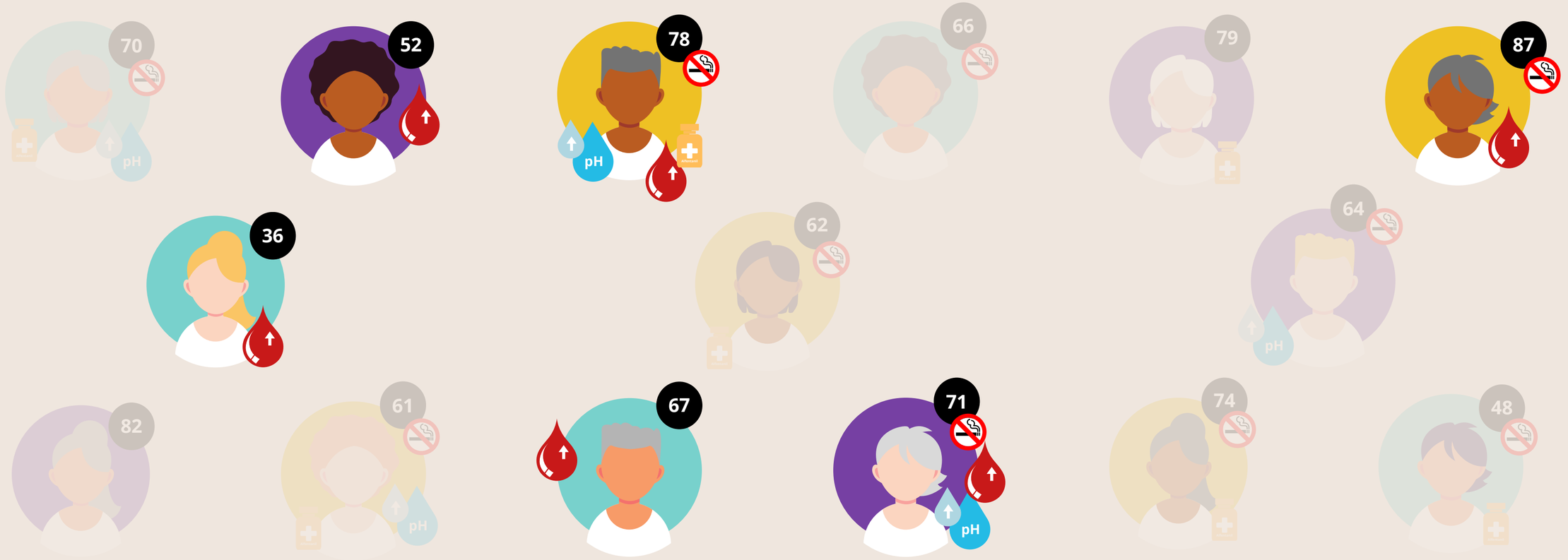
Local Pattern Mining

Exceptional Model Mining for Stratified Medicine



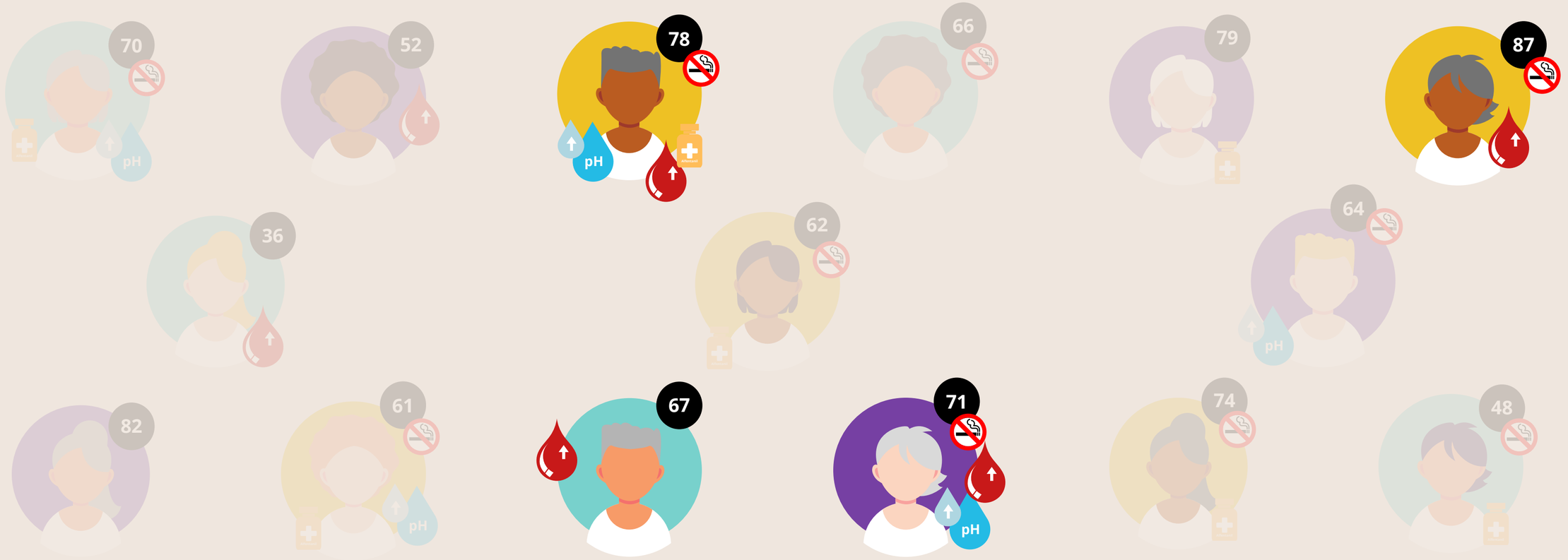
Local Pattern Mining

Exceptional Model Mining for Stratified Medicine



Local Pattern Mining

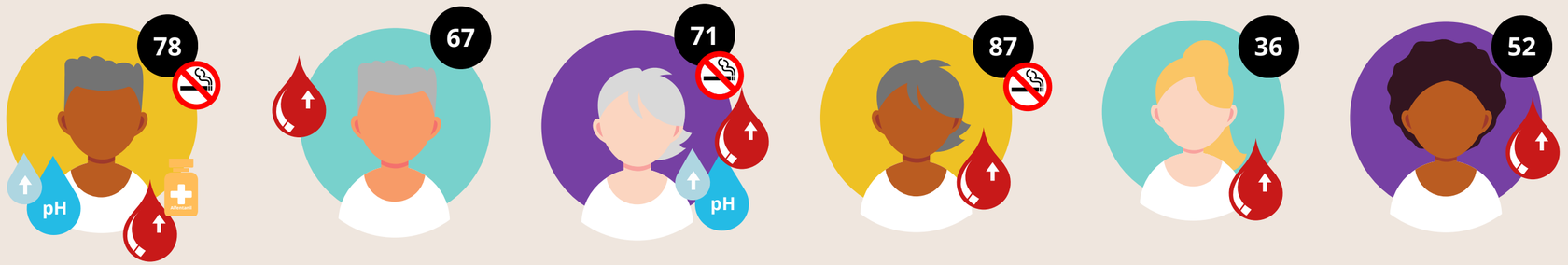
Exceptional Model Mining for Stratified Medicine



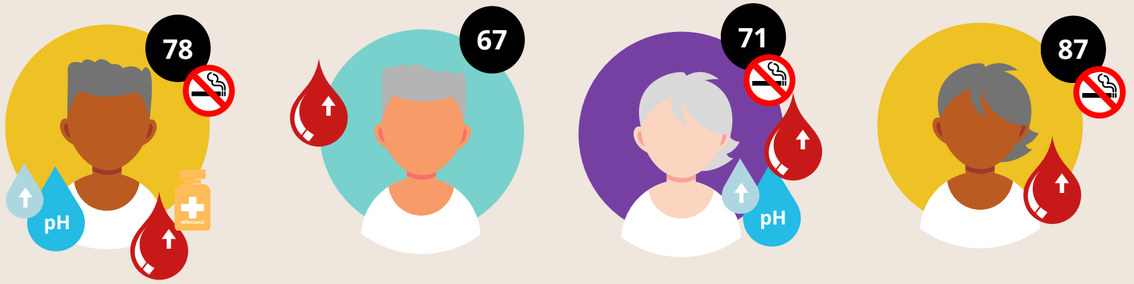
Local Pattern Mining

Exceptional Model Mining for Stratified Medicine

{Blood Loss = high}



{Blood Loss = high AND Age > 60}

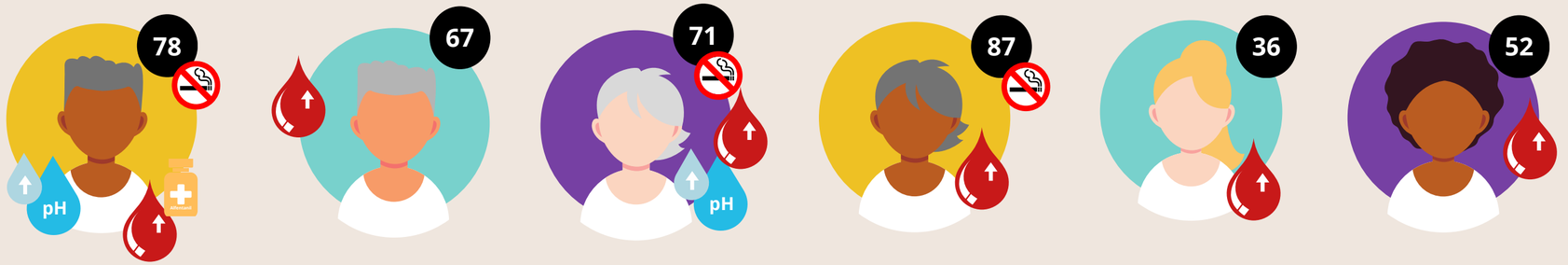


Local Pattern Mining

Exceptional Model Mining for Stratified Medicine

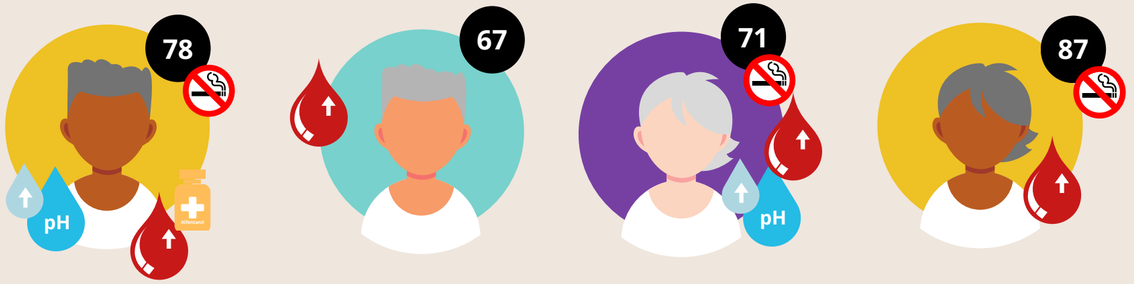
Example quality measure:
% of irregular RR-intervals
in ECG morphology

{Blood Loss = high}



40%

{Blood Loss = high AND Age > 60}

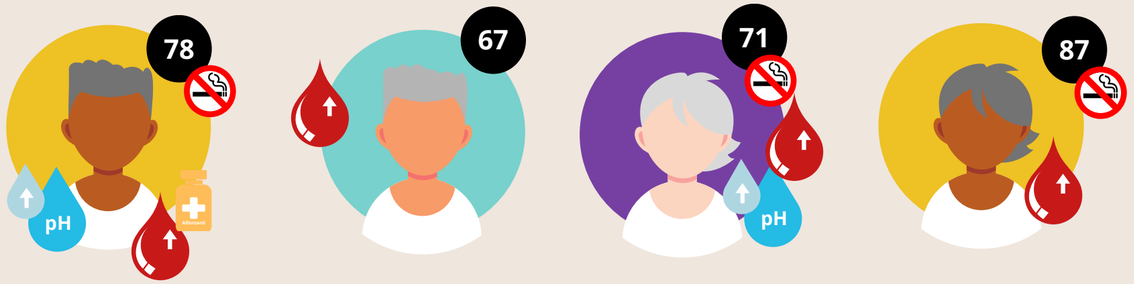


80%

Local Pattern Mining

Exceptional Model Mining for Stratified Medicine

{Blood Loss = high AND Age > 60}



Example quality measure:
% of irregular RR-intervals
in ECG morphology

80%

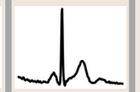
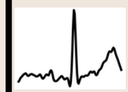
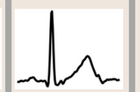
Patients aged over 60 with high blood loss during surgery
have a higher risk of AF development and should get preventive treatment

Exceptional Phenotypes

The EMM Framework for AF Characterization

descriptors | *targets*

patients

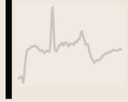
	AGE					...				pH	hb1	hb2	hb...
	56	↑	used	uses	never	...	True	False	False	7.6			...
	43	↓	never	uses	never	...	True	True	True	7.4			...
	89	↑	never	used	never	...	False	False	True	8.6	...	...	...
	77	↑	uses	uses	used	...	True	True	False	7.2	...	...	...
	56	↓	never	never	never	...	False	True	False	6.7			...

Exceptional Phenotypes

Descriptors: Medical Characteristics

descriptors | *targets*

patients

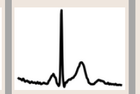
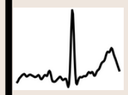
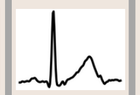
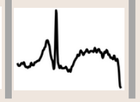
	AGE					...				pH	hb1	hb2	hb...
	56	↑	used	uses	never	...	True	False	False	7.6			...
	43	↓	never	uses	never	...	True	True	True	7.4			...
	89	↑	never	used	never	...	False	False	True	8.6	...	...	...
	77	↑	uses	uses	used	...	True	True	False	7.2	...	...	...
	56	↓	never	never	never	...	False	True	False	6.7			...

Exceptional Phenotypes

Targets: Electrocardiogram Signals of Heartbeats

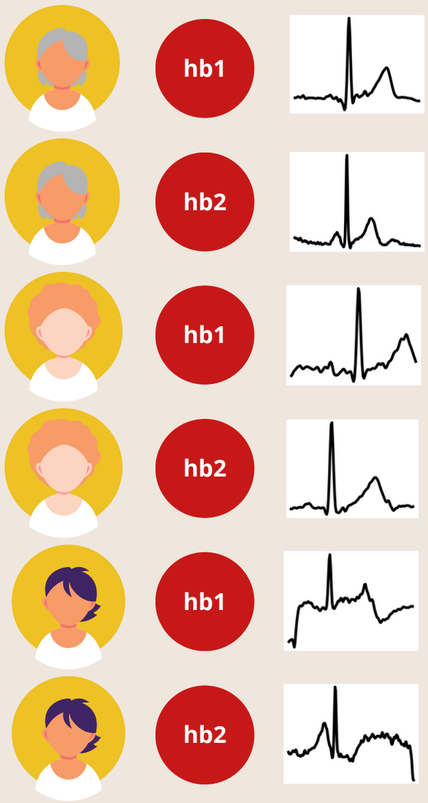
descriptors | *targets*

patients

	AGE					...					hb1	hb2	hb...
	56	↑	used	uses	never	...	True	False	False	7.6			...
	43	↓	never	uses	never	...	True	True	True	7.4			...
	89	↑	never	used	never	...	False	False	True	8.6	...	...	...
	77	↑	uses	uses	used	...	True	True	False	7.2	...	...	...
	56	↓	never	never	never	...	False	True	False	6.7			...

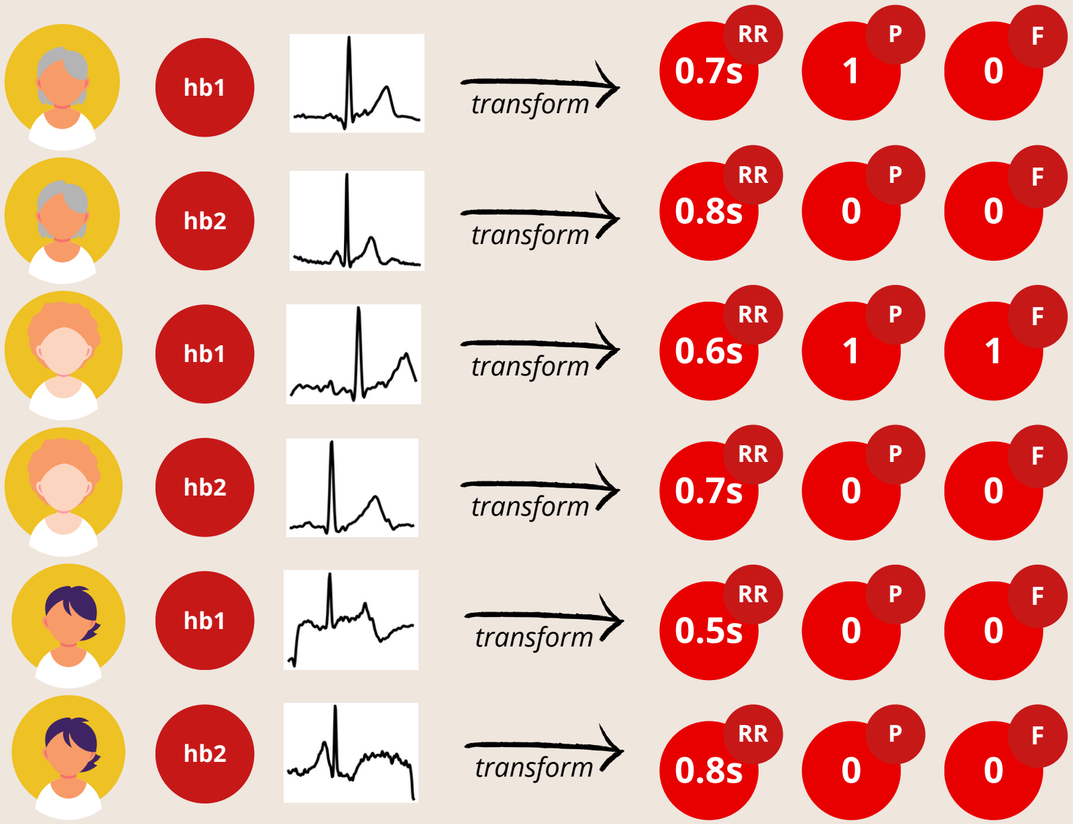
Exceptional Phenotypes

Targets: Electrocardiogram Signals of Heartbeats



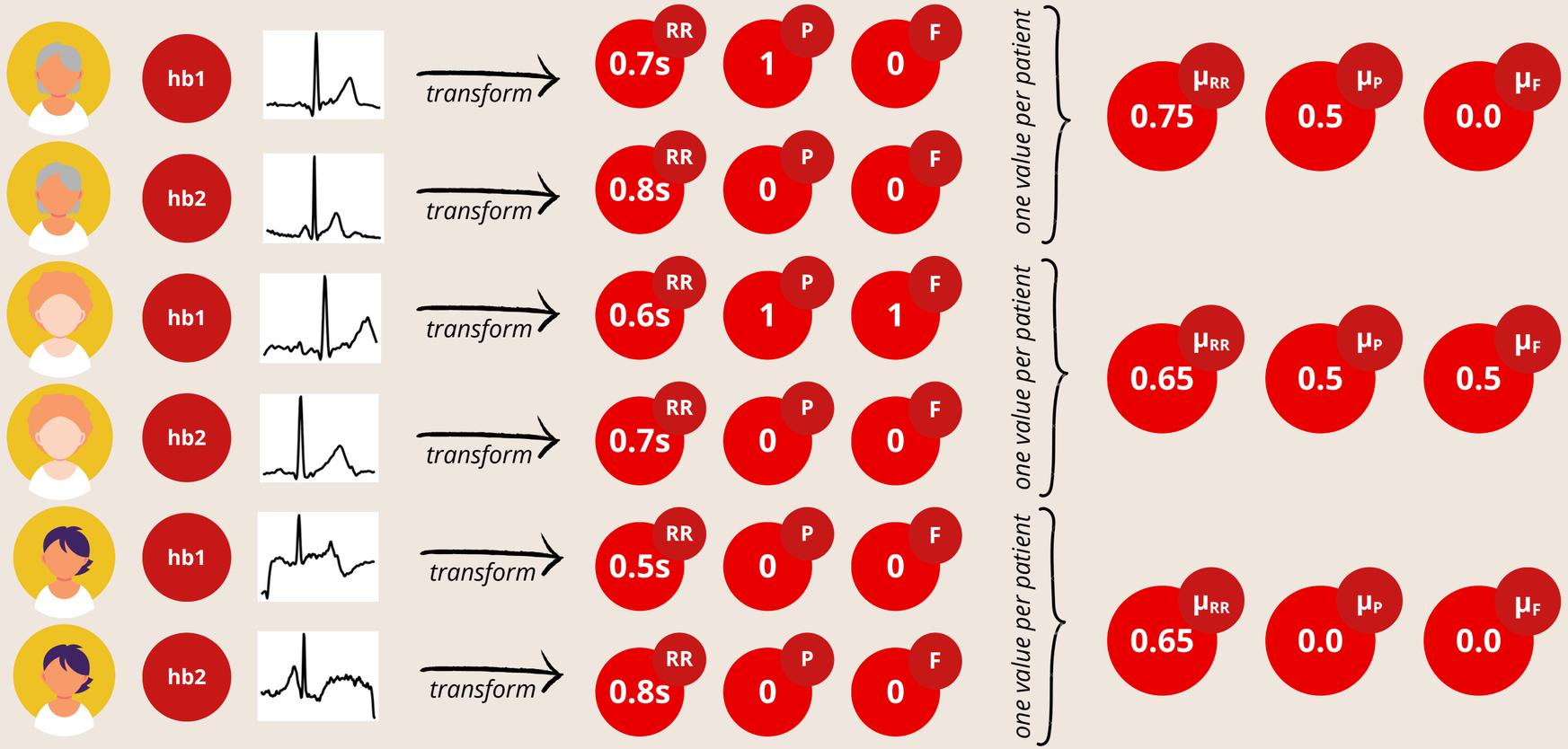
Exceptional Phenotypes

Targets: Electrocardiogram Signals of Heartbeats



Exceptional Phenotypes

Targets: Electrocardiogram Signals of Heartbeats



Exceptional Phenotypes

Quality Measures

Exceptionality factor, including:

- entropy function;
- precision function;
- phenotype function.

$$\varphi(D) = \varphi_{\text{ef}}(D) \cdot \varphi_{\text{pr}}(D) \cdot \varphi_{\text{pheno}}(D)$$

Exceptional Phenotypes

Quality Measures

Phenotype function, as

- the difference between the population's and subgroup's phenotype.

$$\varphi_{\text{pheno}}(D) = \bar{\theta}^{G_D} - \bar{\theta}^{\Omega}$$

Exceptional Phenotypes

Quality Measures

Phenotypes, related to

- *irregular heart rates;*
- absence of P-waves;
- replacement F-waves;
- combinations thereof.

$$\theta_{\mathrm{SDRR}}(p) = \sqrt{\frac{1}{K-2}\sum_{i=1}^{K-1}(RR_i - \overline{RR})^2}$$

$$\theta_{\mathrm{RMSSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2}(\Delta RR_i)^2}$$

$$\theta_{\mathrm{SDSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2}(\Delta RR_i - \overline{\Delta RR})^2}$$

Exceptional Phenotypes

Quality Measures

Phenotypes, related to

- irregular heart rates;
- *absence of P-waves;*
- *replacement F-waves;*
- combinations thereof.

$$\theta_P(D) = \frac{1}{K}\sum_{i=1}^{K} P_i$$

$$\theta_F(D) = \frac{1}{K}\sum_{i=1}^{K} F_i$$

Exceptional Phenotypes

Quality Measures

Phenotypes, related to

- irregular heart rates;
- absence of P-waves;
- replacement F-waves;
- *combinations thereof.*

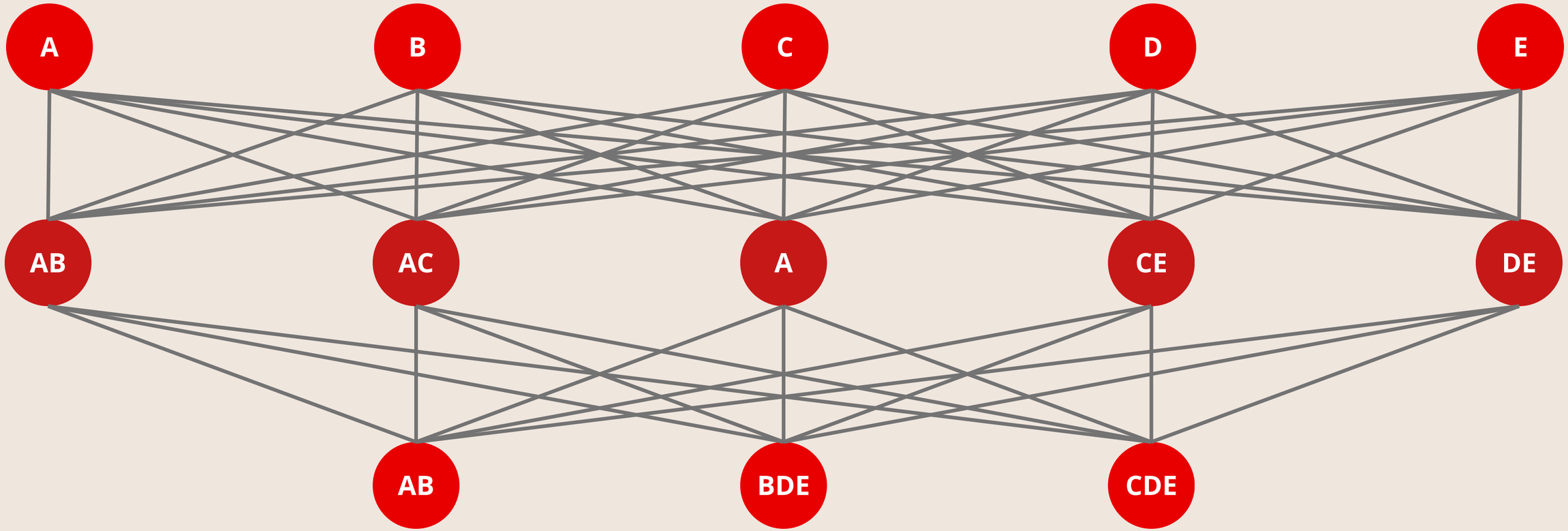
$$\theta^*_{\mathrm{SDSQ}}(p) = \sqrt{\frac{1}{K-2}\sum_{i=1}^{K-1} \mathbf{1}_*(p,i)\cdot(SQ_i - \overline{SQ})^2}$$

$$\theta^*_{\mathrm{RMSSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2} \mathbf{1}_*(p,i)\cdot(\Delta SQ_i)^2}$$

$$\theta^*_{\mathrm{SDSD}}(p) = \sqrt{\frac{1}{K-3}\sum_{i=1}^{K-2} \mathbf{1}_*(p,i)\cdot(\Delta SQ_i - \overline{\Delta SQ})^2}$$

Exceptional Phenotypes

Beam Search for Subgroups

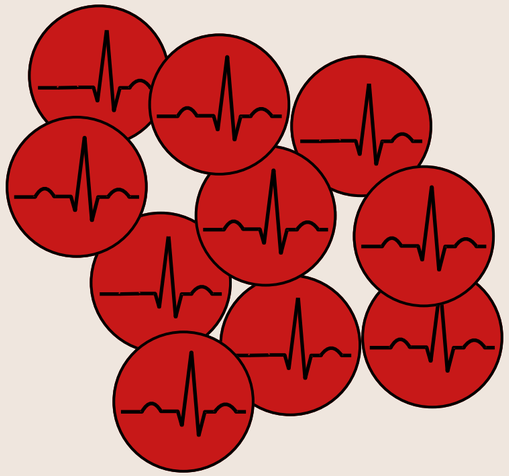


Exceptional Phenotypes

Beam Search for Subgroups

AGE > 70

40% of irregular RR-intervals
in ECG morphology

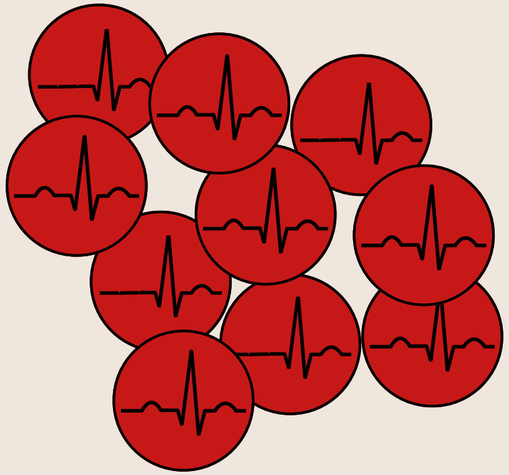


Exceptional Phenotypes

Beam Search for Subgroups

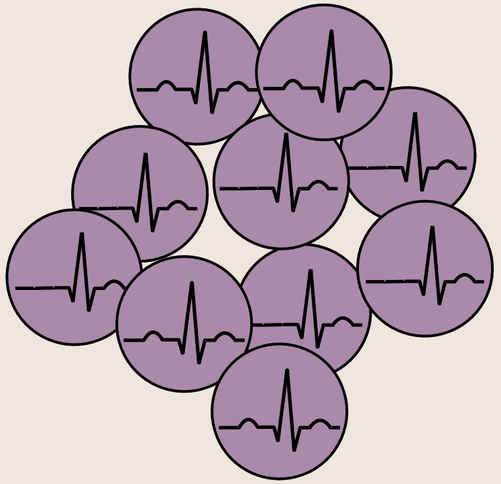
AGE > 70

40% of irregular RR-intervals
in ECG morphology



CHLORIDE > 106

60% of irregular RR-intervals
in ECG morphology

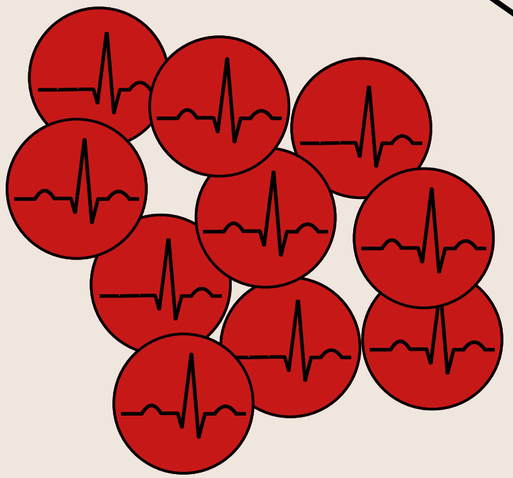


Exceptional Phenotypes

Beam Search for Subgroups

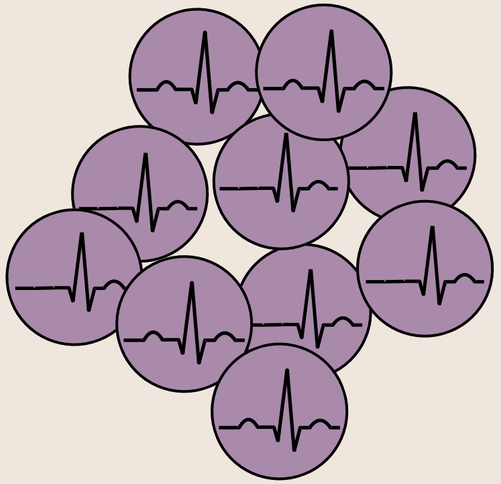
AGE > 70

40% of irregular RR-intervals
in ECG morphology



CHLORIDE > 106

60% of irregular RR-intervals
in ECG morphology



AGE > 70 AND CHLORIDE > 106

80% of irregular RR-intervals
in ECG morphology

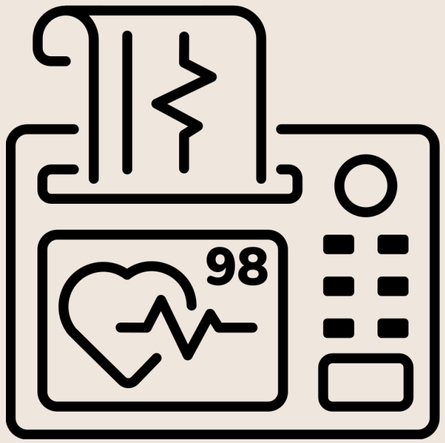
Real-Life Data Experiments

Data from Catharina Hospital

Descriptors

Electronic Health Records

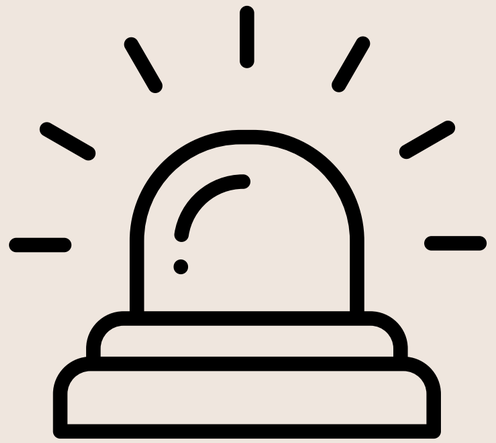
247 aggregated
patient characteristics



Targets

Electrocardiogram Signals

Atrial Fibrillation features
extracted from Lead II



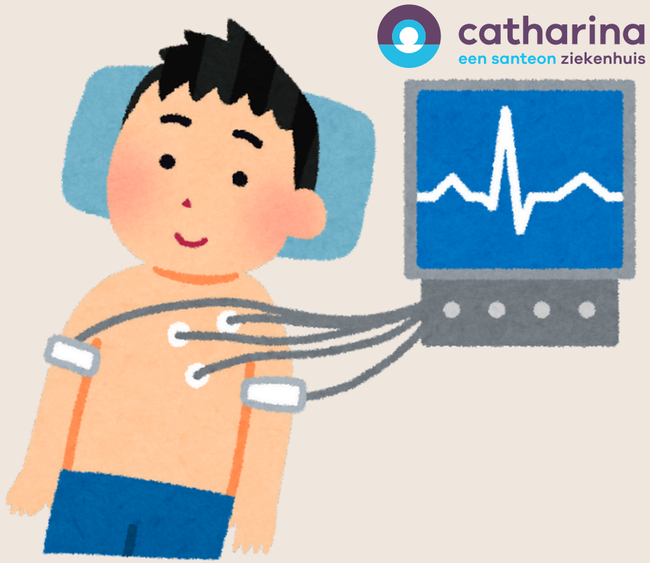
Evaluators

Atrial Fibrillation Complications

Alarm indicators during and
<4 weeks after surgery

Real-Life Data Experiments

Patient Population



230 cardiac patients admitted in Catharina Hospital for cardiac surgery

Resulting Subgroups

Summary Statistics of the Eleven Experiments

Singular Phenotypes

#	phenotype	Δ exceptionality	% AF*
1	SDSD	+37%	71%
2	RMSSD	+51%	71%
3	SDRR	+39%	73%
4	P-wave absence	+66%	71%
5	F-waves	+248%	70%

* AF percentage of full dataset is 37.8%.

Combined Phenotypes

#	phenotype	Δ exceptionality	% AF*
6	SDSD & P-wave	+167%	60%
7	RMSSD & P-wave	+171%	58%
8	SDSQ & P-wave	+65%	57%
9	SDSD & F-waves	+205%	59%
10	RMSSD & F-waves	+171%	58%
11	SDSQ & F-waves	+54%	58%

Resulting Subgroups

Summary Statistics of the Eleven Experiments

Singular Phenotypes

#	phenotype	Δ exceptionality	% AF*
1	SDSD	+37%	**71%**
2	RMSSD	+51%	**71%**
3	SDRR	+39%	**73%**
4	P-wave absence	+66%	**71%**
5	F-waves	**+248%**	**70%**

* AF percentage of full dataset is 37.8%.

Combined Phenotypes

#	phenotype	Δ exceptionality	% AF*
6	SDSD & P-wave	**+167%**	60%
7	RMSSD & P-wave	**+171%**	58%
8	SDSQ & P-wave	+65%	57%
9	SDSD & F-waves	**+205%**	59%
10	RMSSD & F-waves	**+171%**	58%
11	SDSQ & F-waves	+54%	58%

Resulting Subgroups

Summary Statistics of the Eleven Experiments

Singular Phenotypes

#	phenotype	Δ exceptionality	% AF*
1	SDSD	+37%	**71%**
2	RMSSD	+51%	**71%**
3	SDRR	+39%	**73%**
4	P-wave absence	+66%	**71%**
5	F-waves	**+248%**	**70%**

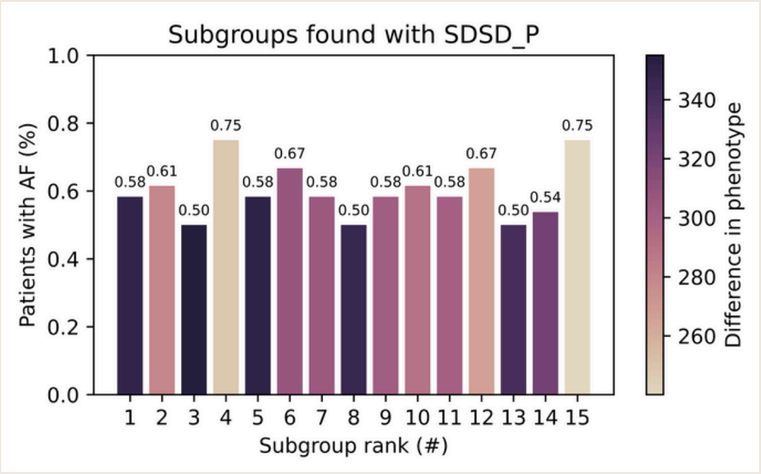
* AF percentage of full dataset is 37.8%.

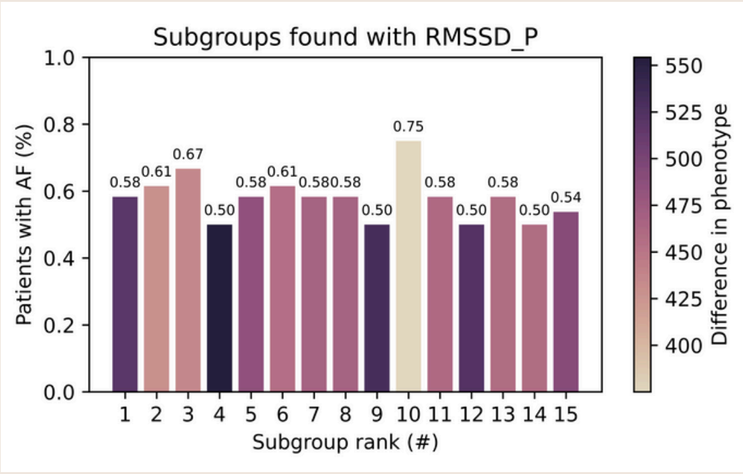
Combined Phenotypes

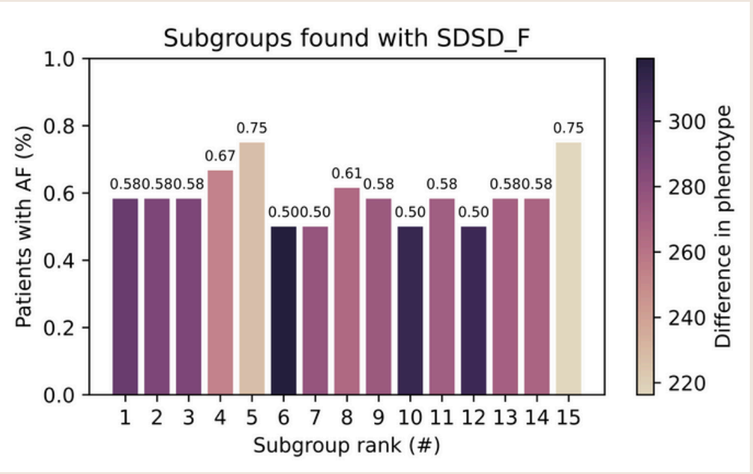
#	phenotype	Δ exceptionality	% AF*
6	**SDSD & P-wave**	**+167%**	60%
7	**RMSSD & P-wave**	**+171%**	58%
8	SDSQ & P-wave	+65%	57%
9	**SDSD & F-waves**	**+205%**	59%
10	RMSSD & F-waves	**+171%**	58%
11	SDSQ & F-waves	+54%	58%

Resulting Subgroups

Evaluation of Three Selected Experiments







The height of the bars indicate the percentage of patients in the subgroup that experienced AF.
The darkness of the bars indicate the exceptionality of the phenotype: the darker the greater the exceptionality.

Resulting Subgroups

Identified Subgroups

a. Patients that are assisted by the heart-lung machine that also have acidic blood;

b. Patients with blood clotting problems that need cefazolin admission;

c. Patients with high chloride levels;

d. Patients with blood clotting problems that need Ringer’s lactate to overcome low blood volume/low blood pressure;

e. Patients that get Alfentanil administered.

Future Work

Medical direction

- Clinical follow-up studies on the found risk factor combinations.
- Repetition of this study at other hospitals.
- Generalization of this study to other cardiac diseases.

Technical direction

- Combining more leads to catch more subtle AF in the ECGs.
- Scaling the method towards a predictive technique.
- Introducing an evaluation metric scaled towards medical applications.

Collaborators

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

catharina
een santeon ziekenhuis

Contact

Lieke van den Biggelaar
PhD Candidate in Data Mining
Eindhoven University of Technology
✉ l.a.j.v.d.biggelaar@tue.nl

Paper

Code